International Journal of Engineering & Technology, 7 (4.10) (2018) 935-939

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper



Detection of Epileptic Seizure Using Wavelet Analysis based Shannon Entropy, Logarithmic Energy Entropy and Support Vector Machine

Vasudha Harlalka[1], Viraj Pradip Puntambekar[2], Kalugotla Raviteja[3], P.Mahalakshmi[4*]

[1,2,3,4]School of Electrical Engineering,
VIT, Vellore, Tamil Nadu, India.
**Corresponding author E-mail: pmahalakshmi@vit.ac.in*

Abstract

Epilepsy is a prevalent condition, mainly affecting the nervous system of the human body. Electroencephalogram (EEG) is used to evaluate and examine the seizures caused due to epilepsy. The issue of low precision and poor comprehensiveness is worked upon using dual tree- complex wavelet transform (DT-CWT), rather than discrete wavelet transform (DWT). Here, Logarithmic energy entropy (LogEn) and Shannon entropy (ShanEn) are taken as input features. These features are fed to Linear Support Vector Machine (L-SVM) Classifier. For LogEn, accuracy of 100% for A-E, 99.34% for AB-E, and 98.67% for AC-E is achieved. While ShanEn combinations give accuracy of 96.67% for AB-E and 95.5% for ABC-E. These results showcase that our methodology is suitable for overcoming the problem and can become an alternate option for clinical diagnosis.

Keywords: *Shannon Entropy (ShanEn), Logarithmic energy entropy (logEn), EEG, epilepsy, dual tree complex wavelet transform (DT-CWT)*

1. Introduction

Epilepsy is a neurological disorder, involving repeated seizures, where involuntary movement of entire or partial body happens. It often includes loss of consciousness, loss of bladder control. According to latest survey by the World Health Organisation [WHO], approximately 50 million worldwide have epilepsy, which makes it a widespread condition [1]. The widely used method for clinical diagnosis of neurological disorders including epilepsy is Electroencephalogram (EEG) [1]. Time-frequency domain and nonlinear analysis methods find prominence over time domain and frequency domain, due to superior results owing to the complex and non-linear nature of EEG signals.

Discrete Wavelet Transform (DWT), a type of time-frequency domain analysis, has been commonly used for algorithmic research and efficiency [2]. This is used for epilepsy detection as a feature extractor which is required for finding accuracy corresponding to the data sets. Though, lack of shift invariance becomes a major drawback for feature extracted from DWT.

Therefore, Dual Tree Complex Wavelet Transform (DT-CWT) methodology which ensures shift invariance as a crucial factor is proposed and dealt with in this paper. [3].

Dual Tree Complex Wavelet Transform (DT-CWT) based feature extraction methods namely ShanEn and LogEn is showcased in this paper for epileptic seizure detection [3]. Thereafter, they are classified using Support Vector Machine (SVM). The entire work and related coding has been performed with help of MATLAB [4,5].

The datasets are divided into five levels from Level 2 to Level 6 by DT-CWT. These datasets- A, B, C, D, E- is used for entropy values calculations. The values of ShanEn and LogEn is fed to Support Vector Machine (SVM) for classification into healthy, interictal and ictal. We have chosen a total of 15 cases based on complexity, practical significance and utility. For each case, data of class 1 and class 2 is divided into two halves. The first half of entropy values is taken as training set while, testing set being the second half. Therefore, cases considered in this paper are: -

Cases	Group 1 and Group 2
1.	Dataset A vs Dataset E.
2.	Dataset B vs Dataset E.
3.	Dataset C vs Dataset E.
4.	Dataset D vs Dataset E.
5.	Dataset A-B vs Dataset E.
6.	Dataset A-C vs Dataset E.
7.	Dataset A-D vs Dataset E.
8.	Dataset B-C vs Dataset E.
9.	Dataset B-D vs Dataset E.
10.	Dataset C-D vs Dataset E.
11.	Dataset A, B, C vs Dataset E.
12.	Dataset A, B, D vs Dataset E.
13.	Dataset A, C, D vs Dataset E.
14.	Dataset B, C, D vs Dataset E.
15.	Dataset A, B, C, D vs Dataset E.

The training dataset and testing dataset are fed into the SVM classifier, through which specificity, sensitivity and accuracy are calculated.

This paper is further organized as: DT-CWT method used (elaborate study on proposed method), Procedure for evaluation, Results, Conclusion and References.

2. Data

The five datasets used in thesis is extracted from the Department of Epileptologie in Bonn which has clinics Germany. This clinical

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

data (A-E) [Table 1] contains 100 EEG signals recorded by a single channel. Each dataset takes 23.6s duration which are selected from uninterrupted multichannel EEG signals. EEG signals in Dataset A and Dataset B consists of healthy individuals in eyes open and eyes closed in relaxed state, respectively. The signals obtained are taken from the surface of an individual's head. Dataset C, D and E signals are of epileptic patients. Both Set C and D are pre-ictal and recorded during seizure free period from different part of the brain. Dataset C is recorded before an epileptic attack onset from opposite hemisphere of the brain which has the hippocampal formation, whereas Dataset D is recorded before an attack from the epileptogenic zone of the scalp. Set E is recorded during ictal phase, when epileptic attack happens to a patient.
The Analog EEG signal is converted to computer understandable Digital signal by a 12-bit ADC converter with a sampling frequency of 173.61Hz. Data acquisition system stores the information. A 128-channel amplifier system is used to record the signals [6]. EEG of normal person and epileptic person is shown in Figures 1a and 1b.

3. Methodology

The EEG dataset A to dataset E is decomposed into 6 sub sub-band signals. This is analyzed by DT-CWT instead of DWT. It is followed by analysis of Shannon (or Information) entropy and Logarithmic energy entropy. Thereafter, categorization into normal vs inter-ictal or ictal is done by feeding the stored values of entropies to Support Vector Machine Classifiers.
Dual tree - complex wavelet transform
There are significant advantages of DT-CWT over DWT making it a more suitable analysis method for EEG signals. Since, EEG signals are non-stationary and non-linear signals.
The property of approximate shift invariance and directional selectivity in two and higher dimensions prove advantageous for EEG signals. The implementation of DT-CWT signal processing method is done by MATLAB.
The disadvantages of DWT technique are- shift variance, aliasing effects, and wavelet oscillations. It represents EEG signals using multiscale properties. A signal x is broken into n sampling points, here high pass filter (h) has H, as high pass DWT coefficient and low pass filter(l) has L, as low pass DWT coefficient.
The disadvantage of DWT is overcome by DT-CWT by implementation of two distinct 2-channel filter banks. Also, random selection of scaling and wavelet filter is not possible in DT-CWT. The wavelet and scaling function produced by wavelet and scaling filters, are approximate Hilbert transform for another tree [3].
In DT-CWT, dual tree of wavelet filters computes complex wavelet parameters into real and imaginary parts. Here, DT-CWT decomposes EEG signals into six level. Entropies- Shannon and Logarithmic energy are computed for the decomposed signals. Thereafter, the extracted entropy values are stored, divided into training dataset and testing dataset. This sets are fed to SVM classifiers to calculate sensitivity, specificity and accuracy. Therefore, wavelet generated from the complex-valued scaling function and trees are approximately analytic.

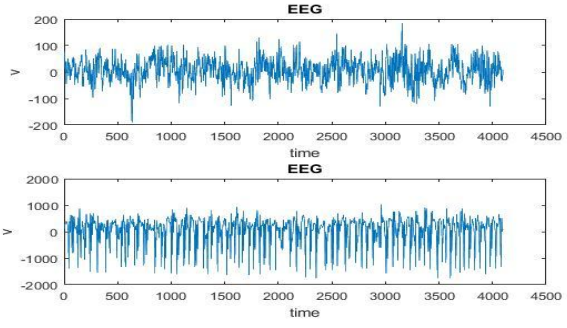


Fig1a, b: EEG of Normal person and Epileptic Person. (Dataset A and E)

Table 1: Clinical Data Sets

Subjects	Set A	Set B	Set C	Set D	Set E
	Healthy Person		Epileptic Patients		
Patient State	Eyes Open	Eyes Closed	Pre-ictal	Pre-ictal	Ictal
Electrode Types	Surface	Surface	Intracranial	Intracranial	Intracranial
Electrode Placement	International 10/20 systems	International 10/20 systems	Within epileptogenic zone	Opposite to epileptogenic zone	Within epileptogenic zone
No. of Samples	100	100	100	100	100
Sampling Points	4096	4096	4096	4096	4096

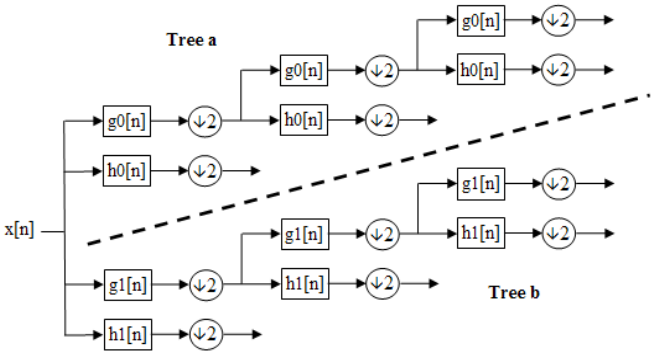


Fig2. Dual tree complex wavelet transform decomposition

3.1. Entropies:

The randomness of signal is measured by entropy. In other words, it reflects quantity of information in the given signal [7]. The entropy is defined by:

$$H(x) = -\sum_{i=0}^{N-1} p_i(x) log_2(p_i(x))$$

where finite discrete random variable, x=[x(0), x(1),....x(N-1)], p(x) denotes probability distribution function, and i indicates one of the discrete states. Hence, entropy is higher when discrete states have equal probability of occurrence.
Shannon or Information Entropy was introduced by Claude Shannon in his paper. In simplistic terms, Shannon Entropy is number of binary digits required to encode a message, which depends on symbols and its frequency.
Equation for Shannon or Information Entropy is:

$$H(x) = -\sum_{0}^{N-1} p_x log_2 p_x$$

where p_x represents the probability.

$$log_2(n) = \frac{log_3(n)}{log_3(2)} \text{ ---(i)}$$

Converting from one logarithmic base to logarithmic base 2 is done by aforementioned formulae.
Similarly, logarithmic energy entropy is given by, [7]

$$H_{LogEn}(x) = -\sum_{i=0}^{N-1} (log_2(p_i(x)))^2$$

3.2. Support Vector Machines (SVM)

Support Vector Machines (SVM) is a supervised learning algorithm, under machine learning. It can be implemented for both classification and regression, but mainly finds usage in classification. SVM can be used for high dimensional feature vectors too [8]. It is a binary classifier model which aims to divide dataset into two classes [9]. This is done by finding a hyperplane using support vectors.

$$w.x_i + b \geq +1, \; if \; y_i = +1$$
$$w.x_i + b \leq +1, \; if \; y_i = -1$$

Hyperplane is a line which divides data into two classes with widest possible gap. Support vectors are data points nearest to hyperplane. Thus, any change in its location can alter the position hyperplane. The dataset close to hyperplane is noise sensitive, thereby hyperplane is kept far from data points. The margins determined by input vectors are the support vectors. A kernel function is applied, if the problem is not linearly separable. Kernel function such as linear, polynomial, radial basis and sigmoidal function are applied to support vectors, depending upon problem. This converts the support vector to linearly separable in transformed space [3].

In this paper, dataset is divided into- training set and testing set. The classifier is fed these sets to evaluate the desired parameters. When tested with unknown or random data points, it improves the methodology and gives better accuracies. Also, calculating average of accuracies achieved by various data makes the method more trustworthy.

3.3. Parameters

The parameters used to obtain the performance of the Support Vector Machine are given below:

Specificity(SPE) [10]

Sensitivity(SEN)[10]

Classification Accuracy(CA)

Their calculations are done as follows:

$$specificity(SPE) = \frac{true\ negatives}{true\ negatives + false\ possitives} * 100\%$$

$$sensitivity(SEN) = \frac{true\ possitives}{true\ possitive + false\ negatives} * 100\%$$

$$classification\ accuracy(CA) = \frac{correct\ classified\ patterns}{total\ patterns} * 100$$

4. Results

4.1 Results Obtained

The complete set of data from A to E, which was obtained from University of Bonn is decomposed using DT-CWT into six levels. However, level 2 to level 6 is useful and further analysed. Figure 2 represents the signal after DT-CWT application. The entropy is calculated for each set of data. It is observed that Set A and Set B have the highest entropy, Set C and Set D have entropies lower than A and B, while Set E has the least entropy values. We have considered a total of 15 cases for analysis, as specified earlier.

The entropy values are divided into two halves- training set and testing set, as class 1 and class 2. These values are stored in excel sheets, fed to the SVM classifiers to evaluate parameters of accuracy, specificity and sensitivity. Upon observation, it was noted that LogEn gave higher accuracy than ShanEn. It was also noted that for lower entropies, LogEn is better than ShanEn. This is done in MATLAB 2015b.

There are 100 samples in each dataset of the Feature Vectors ranging from A to E, where 5 features for observations in each case below.

Case 1-4: For 200 observations, 5 features were taken.

Case 5-10: For 300 observations, 5 features were taken. Case11-14: For 400 observations, 5 features were taken.

Case 15: For 500 observations, 5 features were taken.

In SVM, training and testing data both consisted of 50% of total observations. The SEN, SPE and Classification Accuracy Values are listed below.

Table 2 shows when SVM classifies the feature vector using ShanEn. Table 3shows result using feature vector derived from LogEn.

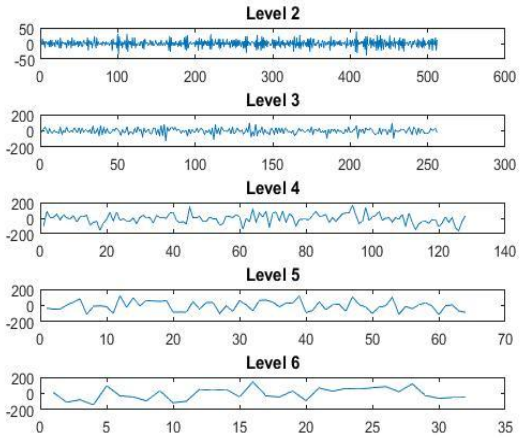


Fig 2a: Dual tree complex wavelet transform on dataset A- Level 2 to 6

Table 2: Results of Support Vector Machine using ShanEn

Sets(Cases)	Elapsed Time	Accuracy	Specificity	Sensitivity
A-E	0.015756	94	89.2857	100
B-E	0.015588	92	86.1978	100
C-E	0.015603	92	86.1978	100
D-E	0.015968	84	75.7675	100
AB-E	0.049994	96.6667	95.2381	100
AC-E	0.015986	96	94.3401	100
AD-E	0.016978	90.6667	90.0233	92.8642
BC-E	0.015688	96	94.3401	100
BD-E	0.022338	90.6667	89.9438	90.8967
CD-E	0.039982	90.6667	90.0233	92.8655
ABC-E	0.016725	95.5000	96.6887	91.8411
ABD-E	0.015734	93.5000	94.2214	91.2222
ACD-E	0.016419	94	94.8052	91.3043
BCD-E	0.015778	93.5000	94.2042	91.2222
ABCD-E	0.015868	95.2000	96.5347	89.6014

Table 3: Results of Support Vector Machine using LogEn

Sets(Cases)	Elapsed Time	Accuracy	Specificity	Sensitivity
A-E	0.015868	100	100	100
B-E	0.065429	98	96.1498	100
C-E	0.019245	98	96.1498	100
D-E	0.024598	97	97.9295	96.0873
AB-E	0.153468	99.3333	99.0099	100
AC-E	0.015763	98.6667	98.0392	100
AD-E	0.015982	97.3333	98.9697	94.2279
BC-E	0.016311	98	98.0201	97.9278
BD-E	0.016300	96	97	94
CD-E	0.022941	96.6667	97.0317	95.9209
ABC-E	0.015960	98.5000	98.6822	97.9278
ABD-E	0.015707	96.5000	98.6415	90.6021
ACD-E	0.016665	98	98.6667	96
BCD-E	0.023193	96	97.9730	90.4146
ABCD-E	0.015670	97.2000	98.9948	90.6560

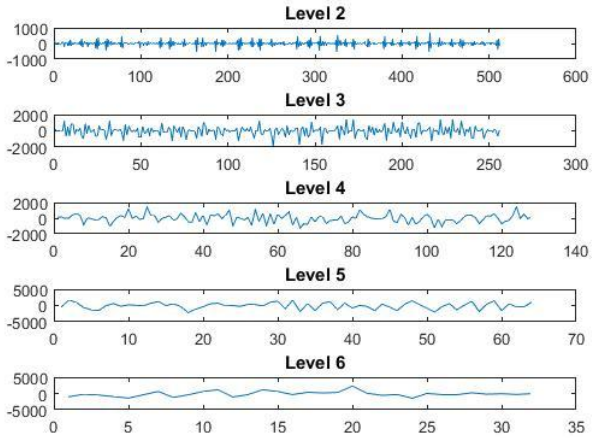


Fig 2b: D Dual tree complex wavelet transform on dataset E- Level 2 to 6.

Table 6.3: Comparison of various methodologies (ascending)

Existing Literature	Case	CA
"Nigam VP1, Graupe D. A neural-network-based detection of epilepsy (2004)"[11]	A-E	97.20
"N. Kannathal et al. Entropies for detection of epilepsy in EEG(2005)" [12]	A-E	92.22
"V. Srinivasan, C. Eswaran, and N. Sriraam, Artificial Neural Network Based Epileptic Detection Using TimeDomain and Frequency Domain Features (2005)" [13]	A-E	99.60
"K. Polat, and S. Güneş, Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform (2007)"[14]	A-E	98.72
"A. Subasi, Signal classification using wavelet feature extraction and a mixture of expert model (2007)" [15]	A-E	95
"T. Tzallas, Epileptic Seizure Detection in EEGs using Time-Frequency Analysis (2009)" [16]	A-E, ABCD-E	100,97.73
"L. Guo et al. Classification of EEG signals using relative wavelet energy and artificial neural networks (2009)" [17]	A-E	95.20
"Hasan Ocak, Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy (2009)"[18]	A-E, ACD-E	99.60, 99.65
"A. Subasi et al. EEG signal classification using PCA, ICA, LDA and support vector machines (2010)"[19]	A-E	99.75(PCA) 99.50(ICA) 100 (LDA)
"Elif Derya Ubeyli, Least squares support vector machine employing model-based methods coefficients for analysis of EEG signals (2010)"[20]	A-E	99.56
"Ali Motie Nasrabadi, et al. Automated detection of epileptic seizures using mixed- methodology: Wavelet-Chaos-KNN Classifier-Mutual Information (2014)"[21]	A-E	99.56
"Yatindra Kumar et al. Epileptic seizure detection using DWT based fuzzy approximate entropy and support vector machine (2014)" [22]	A-E, B-E, C-E, D-E, ACD-E, BCD-E, ABCD-E	100,100 99.6,95.85 98.15 98.22 97.38

4.2 Comparison of Various Methods

The numerous literature on epileptic seizure detection indicate various methodologies studied and analysed by researchers. Table 4, in ascending order of year, compares the classification accuracy of Nigam and Graupe [11], Kannath et al[12], Srinivasan et al[13], Polat and Gunes[14], Subasi[15], Tzallas et al[16], Guo et al[17], Ocak[18], Subasi et al[19], Ubeyli[20], Guo et al[21], and Yatindra Kumar[22]. The ones with similar cases have been considered for clarity [11-22]. In the case A-E accuracy of 100%, same as Yatindra Kumar and Subasi et al, is achieved. Also, in number of cases, our classification accuracy is better than those obtained by earlier work.

5. Conclusion

DT-CWT technique has been implemented to study EEG signal and detect seizures. Shannon entropy and Logarithmic energy entropy are fed as input to Support Vector Machine classifier. It was noted that during epileptic seizures entropy values are lower than normal. This is due to more complex electrophysiological behaviour under normal circumstances for a healthy brain. Hence, the severity of epilepsy is determined by a lower entropy value.
LogEn has been proved to work better with lower entropies as compared to ShanEn. Also, we have achieved 100% accuracy (CA) in case 1 of log entropy, and 96.67% in case 5 of Shannon Entropy. Hence, DT-CWT shows more promising DWT. This has the potential to be used in the medical field to study EEG signals.
In terms of further scope, other pre-processing methods can be considered, more patients dataset can be utilised for studies.

References

[1] WorldHealth Organization, Epilepsy. http://www.who.int/mental_health/neurology/epilepsy/en/index.html
[2] Li M, et al. Automatic epilepsy detection using wavelet-based nonlinear analysis and optimized SVM. Biocybern Biomed Eng (2016)
[3] Dattaprasad Torse, et al. Classification of EEG Signals in a Seizure Detection System Using Dual Tree Complex Wavelet Transform and Least Squares Support Vector Machine(2017)
[4] Mohammad Zavid Parvez and Manoranjan Paul. Seizure Prediction by Analyzing EEG Signal based on Phase Correlation(2015). Conference Paper. DOI:10.1109/ICCITechn.2014.6997315
[5] Suvadeep Bose, B.Tech, NIT Warangal, V Rama, Assistant Professor, NIT Warangal, and Dr. C.B.Rama Rao, Associate Professor, NIT Warangal. EEG signal analysis for Seizure detection using Discrete Wavelet Transform and Random Forest(2017). Computer and Applications (ICCA)
[6] EEG Time Series Data (Department of Epileptology University of Bonn, Germany).http://epileptologie-bonn.de/cms/front_content.php?idcat=193&lang=3&changelang=3
[7] Serap Aydın, et al. Log Energy Entropy-Based EEG Classification with Multilayer Neural Networks in Seizure(2009).
[8] K. Sercan Bayram, et al. Classification of EEG Signals by using Support Vector Machines(2013).
[9] B. Suguna Nanthini and B. Santhi. EEG Signal Analysis for Epileptic Seizure Detection Using Soft Computing Techniques, 2014. World Applied Sciences Journal 29 (Data Mining and Soft Computing Techniques.
[10] N. Wang and M. R. Lyu, "Exploration of instantaneous amplitude and frequency features for epileptic seizure prediction," in Proceedings of the 12th IEEE International Conference on BioInformatics and BioEngineering (BIBE '12), pp. 292–297, November 2012.
[11] Nigam VP1, Graupe D. A neural-network-based detection of epilepsy.,2004. PMID:14977058 .
[12] N. Kannathal , Min Lim Choo , U. Rajendra Acharya , P. K. Sadasivan, Entropies for detection of epilepsy in EEG, Computer Methods and Programs in Biomedicine, v.80 n.3, p.187-194,December2005 doi>10.1016/j.cmpb.2005.06.012].

[13] V. Srinivasan, C. Eswaran, and N. Sriraam, "Artificial Neural Network Based Epileptic Detection Using TimeDomain and Frequency Domain Features", J. Med. Syst., vol.29 (6), 2005, pp. 647-60.
[14] K. Polat, and S. Güneş, "Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform", Appl. Math. Comput., vol. 32 (2), 2007, pp 625-31.
[15] Subasi, "Signal classification using wavelet feature extraction and a mixture of expert model", Exp. Syst. Appl., vol. 32 (4), 2007, pp. 1084-93.
[16] T. Tzallas, "Epileptic Seizure Detection in EEGs using Time-Frequency Analysis," in IEEE transactions on Information Technology, vol. 13 no. 5, September 2009, pp. 703-710
[17] L. Guo, D. Rivero, J. A. Seoane and A. Pazos, "Classification of EEG signals using relative wavelet energy and artificial neural networks",GCE'09, 2009, pp. 12-14
[18] Hasan Ocak, 2009, "Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy", J Expert Systems with Applications 36, 20272036, Elsiever Publications.
[19] Subasi, and M. Ismail Gursoy. "EEG signal classification using PCA, ICA, LDA and support vector machines". Expert Systems with Applications-2010. vol. 37,pp.-8659–8666.
[20] Elif Derya Ubeyli, "Least squares support vector machine employing model-based methods coefficients for analysis of EEG signals" vol 37(1), 2010, pp 233-39
[21] Ali Motie Nasrabadi, et al. Automated detection of epileptic seizures using mixed- methodology: Wavelet-Chaos-KNN Classifier-Mutual Information, 2014.PRZEGLĄD ELEKTROTECHNICZNY (Electrical Review), ISSN 0033-2097, R. 87 NR 4/2011.
[22] Yatindra Kumar , M.L.Dewal, R.S.Anand Department of Electrical Engineering, Indian Institute of Technology,Roorkee,Uttarakhand 247667,India, Epileptic seizure detection using DWT based fuzzy approximate entropy and support vector machine(2014).